Lecture 4
Resemblance Between Relatives

Bruce Walsh. jbwalsh@u.arizona.edu. University of Arizona.
Notes from a short course taught Jan-Feb 2012 at University of Uppsala

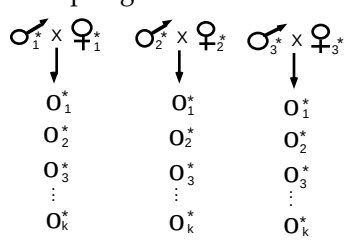
The **heritability** of a trait, a central concept in quantitative genetics, is the proportion of variation among individuals in a population that is due to variation in the additive genetic (i.e., breeding) values of individuals:

$$h^2 = \frac{V_A}{V_P} = \frac{\text{Variance of breeding values}}{\text{Phenotypic Variance}} \tag{4.1}$$

Since an individual's phenotype can be directly scored, the phenotypic variance V_P can be estimated from measurements made directly on the random breeding population.

In contrast, an individual's breeding value cannot be observed directly, but rather must be inferred from the mean value of its offspring (or more generally using the phenotypic values of other known relatives). Thus estimates of V_A require known collections of relatives. The most common situations (which we focus on here) are comparisons between parents and their offspring or comparisons among sibs. We can classify relatives as either ancestral or collateral, and we focus here on designs with just one type of relative. In a more general pedigree, information from both kinds of relatives is present.

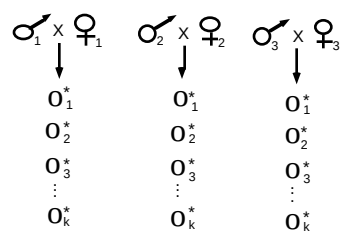
Ancestral relatives: e.g., parent and offspring



* Measure phenotypes of one or both parents, $+k$ offspring of each

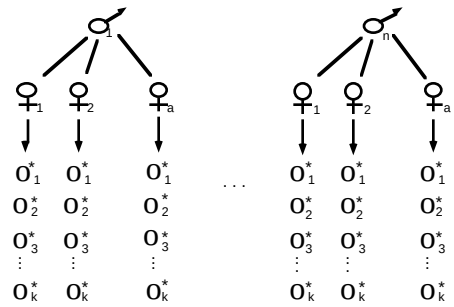
Collateral relatives:

Full Sibs have both parents in common



*Measure k offspring in each family, but not the parents.

Half Sibs have one parent in common



* Measure phenotype of k progeny of each family, but not the parents. Note that if $k > 1$, this design involves both full- (within any column) and half-sibs (between columns from the same sire) and is referred to as a nested half-sib/full-sib design.

Key observation: *The amount of phenotypic resemblance among relatives for the trait provides an indication of the amount of genetic variation for the trait. Further, if trait variation has a significant genetic basis, the closer the relatives, the more similar their appearance.*

Phenotypic Resemblance Between Relatives

We now will use the covariance (and the related measures of correlations and regression slopes) to quantify the phenotypic resemblance between relatives. Quantitative genetics as a field traces back to R. A. Fisher's 1918 paper showing how to use the phenotypic covariances to estimate genetic variances, whereby the phenotypic covariance between relatives is expressed in terms of genetic variances, as we detail below.

1. Parent-offspring regressions

There are three types of parent-offspring regressions: two **single parent - offspring regressions** (plotting offspring mean versus either the trait value in their father P_f or their mother P_m), and the **midparent-offspring regression** (the offspring mean regressed on the mean of their parents, the midparent $MP = (P_f + P_m)/2$).

The slope of the (single) parent-offspring regression is estimated by

$$b_{O|P} = \frac{Cov(O,P)}{Var(P)}, \quad \text{where} \quad Cov(O,P) = \frac{1}{n-1}\left(\sum_{i=1}^{n} O_i\, P_i - n\,\overline{O}\cdot\overline{P}\right)$$

where O_i is the mean trait value in the offspring of parent i and we examine n pairs of parent-offspring. One could compute separate regressions using males (P_m) and females (P_f), although the later potentially includes maternal effect contributions and hence single-parent regressions usually restricted to fathers.

The midparent-offspring regression slope is estimated by

$$b_{O|MP} = \frac{Cov(O,MP)}{Var(MP)}, \quad \text{where} \quad Cov(O,MP) = \frac{1}{n-1}\left(\sum_{i=1}^{n} O_i\, P_{MP,i} - n\,\overline{O}\cdot\overline{MP}\right)$$

where O_i is the mean trait value in the offspring of parents in pair i, where these parents have an average trait value MP_i and we examine n parent-offspring pairs .

Notice that all of the three regressions involve the covariance between parents and their offspring.

2. Collateral relationships: ANOVA

With collateral relatives, the above formulae for the sample covariance is not appropriate, for two reasons. First, there are usually more than two collateral relatives per family. Second, even if families consist of only two relatives, the order of the two is arbitrary — i.e., there is no natural distinction between "X" and "Y", as exists in the case of parents and offspring.

Another way of stating the second point is that collateral relatives belong to the same class or category. In contrast, parents and offspring belong to different classes. The covariance between parents and offspring is an **interclass** (between-class) covariance, while the covariance between collateral relatives is an **intraclass** (within-class) covariance. The analysis of variance (ANOVA), first proposed in Fisher's 1918 paper, is used to estimated intraclass covariances.

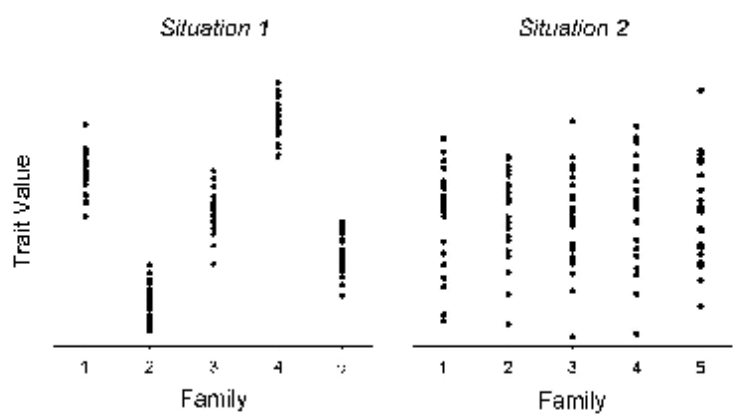
Under the simplest ANOVA framework, we can consider the total variance of a trait to consist of two components: a **between-group** (also called the **among-group**) component (for example, differences in the mean value of different families) and a **within-group** component (the variation in trait value within each family). The total variance is the sum of the between and within group variances,

$$Var(T) = Var(B) + Var(W) \tag{4.2}$$

A key feature of ANOVA is that *the between-group variance equals the within-group covariance*. Thus, the larger the covariance between members of a family, the larger the fraction of total variation that is attributed to differences between family means. To see this result, let $y_{ij} = \mu + b_i + e_{ij}$ be the jth member of group i, where b_i is the group effect and e_{ij} the residual, where $\sigma(e_{ik}, e_{ik}) = 0$. The covariance between two members of group i becomes

$$\sigma(y_{ij}, y_{ik}) = \sigma(\mu + b_i + e_{ij}, \mu + b_i + e_{ik}) = \sigma(b_i, b_i) = \sigma^2(b)$$

the between-group variance (the variance in the group effects).
To see this point, consider the following extreme patterns of phenotypes in full sib families:



Situation 1

Here the between group variance $Var(B) = 2.5$, and the within-group variance $Var(W) = 0.2$. This gives a total phenotypic variance of $V_P = Var(T) = Var(B) + Var(W) = 2.7$. Here:

- members of a family resemble each other more closely than they do members of other families
- there are large differences in average phenotype between families

The resulting intraclass correlation t is

$$t = \frac{Cov(\text{full sibs})}{V_P} = \frac{Var(B)}{V_P} = 0.93$$

where we have used the ANOVA identity that the between-group variance equals the within-group covariance (here, the covariance between full sibs). Since elements of the same class are full-sibs, this is often denoted by t_{FS} to distinguish it from other intraclass correlations.

Situation 2

Suppose the total (phenotypic) variance is the same as in situation 1, with $Var(T) = V_P = 2.7$. However, suppose there is no between-group variance ($Var(B) = 0$), implying that $Var(W) = 2.7$ and the intraclass correlation is $t = 0$. Here:

- members of a family resemble each other no more than they do members of other families
- there are no significant differences in average phenotype between families
- phenotypic resemblance is low, so genetic variation is low

Note that phenotypic resemblance among relatives can equivalently be consider as a measure of the *similiary* among a group of relatives for the phenotype of a quantitative trait (the covariance of family members), or the *difference* in phenotype between different families (the between-group variance), as Cov(Within a group) = Var(Between group means).

Causes of Phenotypic Covariance Among Relatives

Relatives resemble each other for quantitative traits more than they do unrelated members of the population for two potential reasons:

- relatives share genes. The closer the relationship, the higher the proportion of shared genes
- relatives may share similar environment

The Genetic Covariance Between Relatives

The Genetic Covariance $Cov(G_x, G_y)$ = covariance of the genotypic values (G_x, G_y) of individuals x and y.

We will first show how the genetic covariances between parent and offspring, full sibs, and half sibs depend on the genetic variances V_A and V_D. We will then discuss how the covariances are estimated in practice.

Genetic covariances arise because two related individuals are more likely to share alleles than are two unrelated individuals. Sharing alleles means having alleles that are **identical by descent** (IBD): namely that both copies of an allele can be traced back to a single copy in a recent common ancestor. Alleles can also be **identical in state** but not identical by descent.

For example, consider the offspring of two parents and label the four allelic copies in the parents by 1 - 4, independent of whether or not any are identical in state.

$$\text{Parents: } A_1A_2 \times A_3A_4$$

$$\text{Offspring: } o_1 = A_1A_3 \quad o_2 = A_1A_4 \quad o_3 = A_1A_3 \quad o_4 = A_2A_4$$

Here, o_1 and o_2 share one allele IBD, o_1 and o_3 share two alleles IBD, o_1 and o_4 share no alleles IBD.

1. Offspring and one parent

What is the covariance of genotypic values of an offspring (G_o) and its parent (G_p)? Denoting the two parental alleles at a given locus by A_1A_2, since a parent and its offspring share *exactly* one allele, one allele in the offspring came from the parent (say A_1), while the other offspring allele (denoted A_3) came from the other parent. To consider the genetic contributions from a parent to its offspring, write the genotypic value of the parent as $G_p = A + D$. We can further decompose this by considering the contribution from each parental allele to the overall breeding value, with $A = \alpha_1 + \alpha_2$, and we can write the genotypic value of the parent as $G_p = \alpha_1 + \alpha_2 + \delta_{12}$ where δ_{12} denotes the dominance deviation for an A_1A_2 genotype. Likewise, the genotypic value of its offspring is $G_o = \alpha_1 + \alpha_3 + \delta_{13}$, giving

$$Cov(G_o, G_p) = Cov(\alpha_1 + \alpha_2 + \delta_{12}, \alpha_1 + \alpha_3 + \delta_{13}) \tag{4.3}$$

We can use the rules of covariances to expand this covariance between two sums into nine individual covariance terms,

$$\begin{aligned} Cov(G_o, G_p) = & Cov(\alpha_1, \alpha_1) + Cov(\alpha_1, \alpha_3) + Cov(\alpha_1, \delta_{13}) \\ & + Cov(\alpha_2, \alpha_1) + Cov(\alpha_2, \alpha_3) + Cov(\alpha_2, \delta_{13}) \\ & + Cov(\delta_{12}, \alpha_1) + Cov(\delta_{12}, \alpha_3) + Cov(\delta_{12}, \delta_{13}) \end{aligned}$$

By the way have (intentionally) constructed α and δ, they are uncorrelated. Further,

$$Cov(\alpha_x, \alpha_y) = \begin{cases} 0 & \text{if } x \neq y, \quad \text{i.e., not IBD} \\ Var(A)/2 & \text{if } x = y, \quad \text{i.e., IBD} \end{cases} \tag{4.4}$$

The last identity follows since $Var(A) = Var(\alpha_1 + \alpha_2) = 2Var(\alpha_1)$, so that

$$Var(\alpha_1) = Cov(\alpha_1, \alpha_1) = Var(A)/2 \tag{4.5}$$

Hence, when individuals share one allele IBD, they share half the additive genetic variance. Likewise,

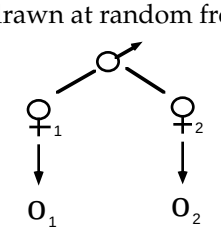
$$Cov(\delta_{xy}, \delta_{wz}) = \begin{cases} 0 & \text{if } xy \neq wz, \quad \text{i.e., both alleles are not IBD} \\ Var(D) & \text{if } xy = wz, \quad \text{both alleles are IBD} \end{cases} \tag{4.6}$$

Two individuals only share the dominance variance when they share both alleles. Using the above identities (Equations 4.4, 4.6), eight of the above nine covariances are zero, leaving

$$Cov(G_o, G_p) = Cov(\alpha_1, \alpha_1) = Var(A)/2$$

2. Half-sibs

Here, one parent is shared, the other is drawn at random from the population;



The genetic covariance between half-sibs is the covariance of the genetic values between o_1 and o_2.

To compute this, consider a single locus. First note that o_1 and o_2 share either one allele IBD (from the father) or no alleles IBD (since the mothers are assumed unrelated, these sibs cannot share both alleles IBD as they share no maternal alleles IBD). The probability that o_1 and o_2 both receive the same allele from the male is one-half (because whichever allele the male passes to o_1, the probability that he passes the same allele to o_2 is one-half). In this case, the two offspring have one allele IBD, and the contribution to the genetic covariance when this occurs is $Cov(\alpha_1, \alpha_1) = Var(A)/2$. When o_1 and o_2 share no alleles IBD, they have no genetic covariance.

Summarizing:

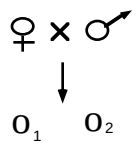
Case	Probability	Contribution
o_1 and o_2 have 0 alleles IBD	$1/2$	0
o_1 and o_2 have 1 allele IBD	$1/2$	$Var(A)/2$

giving the genetic covariance between half sibs as

$$Cov(G_{o_1}, G_{o_2}) = Var(A)/4$$

3. Full-Sibs.

Both parents are in common,



What is the covariance of genotypic values of two full sibs?

As illustrated previously, three cases are possible when considering pairs of full sibs: they can share either 0, 1, or 2 alleles IBD. Applying the same approach as for half sibs, if we can compute: 1) the probability of each case; and 2) the contribution to the genetic covariance for each case.

Each full sib receives one paternal and one maternal allele. The probability that each sib receives the same paternal allele is $1/2$, which is also the probability each sib receives the same maternal allele. Hence,

$$\Pr(2 \text{ alleles IBD}) = \Pr(\text{ paternal allele IBD})\ \Pr(\text{ maternal allele IBD}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\Pr(0 \text{ alleles IBD}) = \Pr(\text{ paternal allele not IBD})\ \Pr(\text{ maternal allele not IBD}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\Pr(1 \text{ allele IBD}) = 1 - \Pr(2 \text{ alleles IBD}) - \Pr(0 \text{ alleles IBD}) = \frac{1}{2}$$

We saw above that when two relatives share one allele IBD, the contribution to the genetic covariance is $Var(A)/2$. When two relatives share both alleles IBD, each has the same genotype at the locus being considered, and the contribution is

$$Cov(\alpha_1 + \alpha_2 + \delta_{12}, \alpha_1 + \alpha_2 + \delta_{12}) = Var(\alpha_1 + \alpha_2 + \delta_{12}) = Var(A) + Var(D)$$

Putting these results together gives

Case	Probability	Contribution
o_1 and o_2 have 0 alleles IBD	$1/4$	0
o_1 and o_2 have 1 allele IBD	$1/2$	$Var(A)/2$
o_1 and o_2 have 2 allele IBD	$1/4$	$Var(A) + Var(D)$

This results in a genetic covariance between full sibs of

$$Cov(G_{o_1}, G_{o_2}) = \frac{1}{2}\frac{Var(A)}{2} + \frac{1}{4}(Var(A) + Var(D)) = \frac{Var(A)}{2} + \frac{Var(D)}{4}$$

Environmental Causes of Relationship Between Relatives

Shared environmental effects (such as a common maternal environment) also contribute to the covariance between relatives, and care must be taken to distinguish these environmental covariances from genetic covariances.

If members of a family are reared together they share a common environmental value, E_c. If the common environmental circumstances are different for each family, the variance due to common environmental effects, V_{Ec}, causes greater similarity among members of a family, and greater differences among families, than would be expected from the proportion of genes they share. Thus, V_{Ec} inflates the phenotypic covariance of sibs over what is expected from their genotypic covariance.

Just as we decomposed the total genotypic value into components, some shared, others not transmitted between relatives, we can do the same for environmental effects. In particular, we can write the total environmental effect E as the sum of a common environmental effect shared by the relatives E_c, a general environmental effect E_g, and a specific environmental effect E_s. Hence, we can write $E = E_c + E_g + E_s$, partitioning the environmental variance as

$$V_E = V_{Ec} + V_{Eg} + V_{Es} \tag{4.7}$$

We can further consider different possible sources of the common environmental effect E_c:

- E_{cS} or E_{cL}: Shared effects due to sharing the space/location (different farms, cages)
- E_{cT}: Temporal (changes in climactic or nutritional conditions over time)
- E_{cM} Maternal (pre- and post-natal nutrition)

Thus, we can partition the environmental variance as

$$\begin{aligned} V_E &= V_{Ec} + V_{Eg} + V_{Es} \\ &= V_{EcS} + V_{EcT} + V_{EcM} + V_{Ec} + V_{Eg} + V_{Es} \end{aligned} \tag{4.8}$$

Common environment effects mainly contribute to resemblance of sibs, but maternal environment effects can contribute to resemblance between mother and offspring as well.

V_{EcS} and V_{EcT} can be eliminated, or estimated, by using the correct experimental design, but it is very difficult (except by cross-fostering) to eliminate or estimate V_{EcM} from the covariance of full sibs. Further, cross-fostering only removes post-natal (past birth) maternal effects, it does not remove shared pre-natal maternal effects.

Complex Relationships in Pedigrees: Coefficients of Coancestry

Much of the analysis in animal breeding occurs with pedigree data, where relationships can be increasingly complex (i.e., inbred relatives). We conclude by introducing the machinery to quantify such relationships.

Coefficients of Coancestry and Inbreeding

Suppose that single alleles are drawn randomly from individuals x and y. The probability that these two alleles are identical by descent, Θ_{xy}, is called the **coefficient of coancestry.** Another way to look at the problem is to consider a hypothetical offspring (z) of x and y. By the above definition, Θ_{xy} is the probability that the two genes at a locus in individual z are identical by descent. The latter quantity is Wright's (1922) **inbreeding coefficient**, f_z. Thus, an individual's inbreeding coefficient is equivalent to its parents' coefficient of coancestry, $f_z = \Theta_{xy}$.

We now proceed by example to demonstrate how estimates of Θ_{xy} are derived. The first problem to be tackled is the coefficient of coancestry of an individual with itself, Θ_{xx}. This may seem like a nonsensical task. However, we will soon see that Θ_{xx} is an essential element of all coancestry estimates. Denote the two genes carried by individual x as A_1 and A_2, and then randomly draw a gene from the locus, replace it, and randomly draw another. Θ_{xx} is the probability that the two genes drawn are identical by descent. There are four ways, each with probability $1/4$, in which the genes can be drawn: A_1 both times, A_1 first and A_2 second, A_2 first and A_1 second, and A_2 both times. If two A_1 genes are drawn, they must be identical by descent since they are copies of the same gene. The same applies to a draw of two A_2 genes. Thus, provided that genes A_1 and A_2 are not identical to each other by descent, then Θ_{xx} is simply $(1/4)(1) + (1/4)(1) = 1/2$. We should, however, recognize the possibility that individual x is inbred, in which case the probability that the gene A_1 is identical by descent with the gene A_2 is f_x. Thus, a general expression for the coefficient of coancestry of an individual with itself is

$$\Theta_{xx} = \frac{1}{4}(1 + f_x + f_x + 1) = \frac{1}{2}(1 + f_x) \tag{4.9}$$

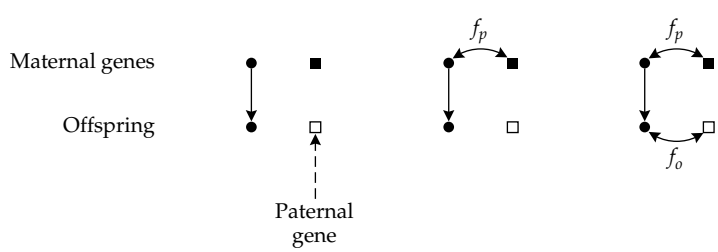


Figure 4.1 The identity of genes by descent for a parent and offspring. Circles and squares represent, respectively, maternally and paternally derived genes. **Left:** The mother is not inbred and her mate is not a relative (so the offspring is not inbred). **Center:** The mother is inbred but unrelated to her mate. **Right:** In addition to the mother being inbred, she is related to her mate, so that her offspring is also inbred.

A slightly more complicated situation arises in calculating the coefficient of coancestry between a parent and its offspring. In order to simplify the discussion, we will call the parent (p) of interest the mother, but the same results apply to fathers provided the locus is autosomal. We first consider the situation in which neither the mother nor her offspring (o) are inbred, i.e., the mother's parents are unrelated, and she is unrelated to her mate. In that case, of the four ways in which single genes can be drawn from the mother and the child, only one involves a pair that is identical by descent (Figure 4.1, left). Therefore, $\Theta_{po} = 1/4$. Suppose, however, that the mother is inbred (Figure 4.1, center), so that the probability that both of her alleles are identical by descent is f_p. This is the same as the probability that the maternal gene inherited by the offspring is identical by descent with the maternal gene not inherited. The probability of drawing such a gene combination is $1/4$. Therefore, inbreeding in the parent inflates Θ_{po} to $(1 + f_p)/4$. With complete inbreeding $(f_p = 1)$, both parental alleles are identical by descent, increasing Θ_{po} to $1/2$. Finally, we allow for

the possibility that the parents of o are related, so that the offspring is inbred with coefficient f_o (Figure 4.1, right). It is now necessary to consider the implications of drawing a paternally derived gene from the offspring, the probability of which is $1/2$. Since f_o is equivalent to the probability that maternally and paternally derived genes are identical by descent, the additional parent-offspring identity induced by inbreeding is $f_o/2$. In summary, the most general expression for the coefficient of coancestry for a parent and offspring is

$$\Theta_{po} = \frac{1}{4}(1 + f_p + 2f_o) \tag{4.10}$$

Often in the literature, Θ_{po} is simply considered to be $1/4$. It should now be clear that this implicitly assumes the absence of matings between relatives.

We now move on to the coefficient of coancestry of two individuals that share the same father and mother (full sibs). We assume a species with separate sexes so that the mother and father are different individuals, and we again start with the simplest situation, progressively allowing the parents to be inbred and/or related (Figure 4.2). For the analysis of full sibs as well as more complicated degrees of relationship, the method of path analysis (Lynch and Walsh Appendix 2) provides a useful tool. The elements in Figure 4.2 no longer represent gametes (as in Figure 4.1) but individuals.

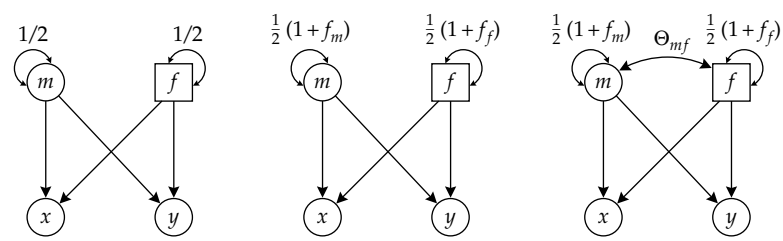


Figure 4.2 Path diagrams for analyzing the probability that random genes from two full sibs are identical by descent. The path coefficients along single-headed arrows are always equal to $1/2$. **Left:** The parents, m and f, are neither related nor inbred. **Center:** The parents are unrelated, but inbred. **Right:** In addition to being inbred, the parents are related with coefficient of coancestry Θ_{mf}.

Let m represent the mother, f the father, and x and y their two offspring. When the parents are neither inbred nor related, there are two paths by which the same gene can be passed to both x and y: $x \leftarrow m \rightarrow y$ and $x \leftarrow f \rightarrow y$. Since both paths have identical consequences, we will simply consider the first of them. First, we note that the probability that both x and y receive the same maternal gene is $1/2$. This is the coefficient of coancestry of the (noninbred) mother with herself, Θ_{mm}, and is represented by the double-headed arrow in the figure. Second, we note that the probability of randomly drawing a maternal gene from individual x is $1/2$, and that the same is true for individual y. Thus, the probability of drawing two maternal genes, identical by descent, one from x and the other from y, is $\Theta_{mm}/4 = 1/8$. Adding the same contribution from the paternal path, $x \leftarrow f \rightarrow y$, we obtain the coefficient of coancestry $\Theta_{xy} = 1/4$. Path analysis provides a simple way to obtain this result. First, set the path coefficients on all of the single-headed arrows in Figure 4.2 equal to $1/2$. Then, note that the contribution of a path to a correlation between two variables is equal to the product of the path coefficients and the correlation coefficient associated with the common factor (in this case, Θ_{mm} or $\Theta_{ff} = 1/2$).

We now allow for the possibility that the parents are inbred with inbreeding coefficients f_m and f_f, a condition that inflates the coefficient of coancestry of an individual with itself. This is the only necessary change for the path diagram in Figure 4.2 (center). There are still only two paths that lead to genes identical by descent in x and y, and their sum is

$$\Theta_{xy} = \frac{1}{4}\left(\Theta_{mm} + \Theta_{ff}\right) = \frac{1}{4}\left(\frac{1+f_m}{2} + \frac{1+f_f}{2}\right) = \frac{1}{8}\left(2 + f_m + f_f\right) \tag{4.11a}$$

Finally, we allow for the possibility that m and f are related, such that the probability of drawing two genes (one from each of them) that are identical by descent is Θ_{mf}. It is then necessary to consider two additional paths between x and y: $x \leftarrow m \leftrightarrow f \rightarrow y$ and $x \leftarrow f \leftrightarrow m \rightarrow y$ (Figure 4.2, right). Again taking the coefficients on the single-headed arrows to be $1/2$, it can be seen that each of these two new paths makes a contribution $\Theta_{mf}/4$ to Θ_{xy}. Adding these to our previous result, we obtain a general expression for the coefficient of coancestry of full sibs,

$$\Theta_{xy} = \frac{1}{8}\left(2 + f_m + f_f + 4\Theta_{mf}\right) \tag{4.11b}$$

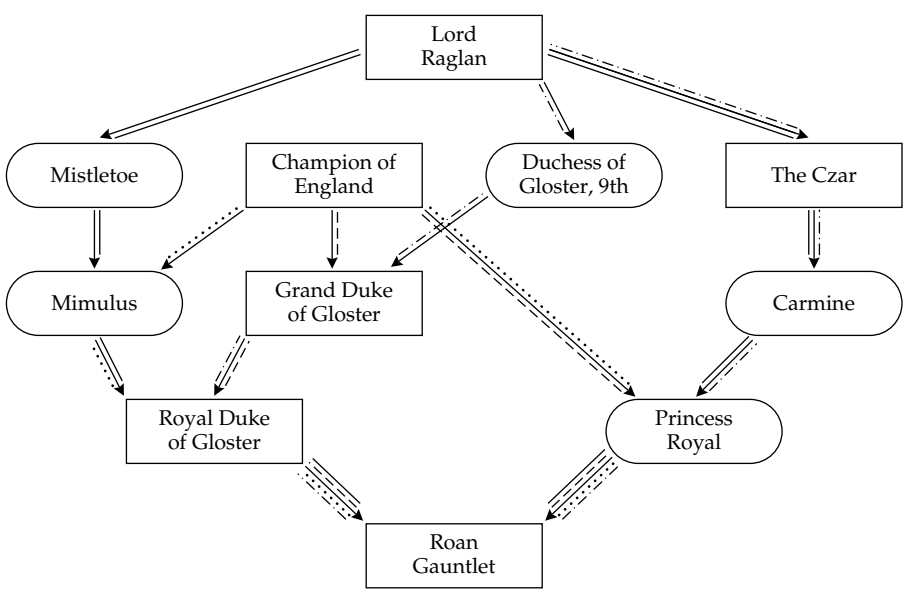
which reduces to $\Theta_{xy} = 1/4$ under random mating.

The preceding techniques are extended readily to more distant relationships and more complicated schemes of relatedness. The coefficient of coancestry is always the sum of a series of two types of paths between x and y. The first type of path leads from a single common ancestor to the two individuals of interest, while the second type passes through two remote ancestors that are related to each other. Neither type of path is allowed to pass through the same ancestor more than once. This procedure is summarized by the following equation

$$\Theta_{xy} = \sum_i \Theta_{ii}\left(\frac{1}{2}\right)^{n_i - 1} + \sum_j \sum_{j \neq k} \Theta_{jk}\left(\frac{1}{2}\right)^{n_{jk} - 2} \tag{4.12}$$

where n_i is the number of individuals (including x and y) in the path leading from common ancestor i, and n_{jk} is the number of individuals (including x and y) on the path leading from two different but related ancestors, j and k.

Example 4.1. One of the first pedigrees to which Wright (1922) applied his theory of inbreeding is that of Roan Gauntlet, an English bull. In the following figure, rectangles and ovals refer to bulls and cows, respectively.



We wish to compute the coefficient of coancestry of the Royal Duke of Gloster and Princess Royal. This is the same as the inbreeding coefficient of their son, Roan Gauntlet. The four possible paths by which alleles identical by descent can be inherited by the Royal Duke and Princess Royal are indicated by the coded lines adjacent to the arrows in the pedigree. Two of these paths contain four individuals and two contain seven. Thus, assuming that the remote ancestors, Lord Raglan and Champion of England, are not inbred (so that for both, $\Theta_{ii} = 1/2$), the coefficient of coancestry of the Royal Duke and Princess Royal is $[2(1/2)^4 + 2(1/2)^7] = 0.141$. This is a slightly closer relationship than that for half sibs (for which $\Theta = 0.125$). Relative to the base population, the alleles at 14% of the autosomal loci in the offspring, Roan Gauntlet, are expected to be identical by descent.

The Coefficient of Fraternity

Up to now we have been considering the identity of single genes by descent. Another useful measure is the probability that single-locus genotypes (both genes) of two individuals are identical by descent. The formulation of such a measure, which we denote as Δ_{xy}, is attributable to Cotterman (1954) and was called the **coefficient of fraternity** by Trustrum (1961). The problem is set out in Figure 4.3. Here we denote the mothers of individuals x and y as m_x and m_y, and the fathers as f_x and f_y. The coefficients of coancestry $\Theta_{m_x m_y}$, $\Theta_{m_x f_y}$, $\Theta_{f_x m_y}$, and $\Theta_{f_x f_y}$ provide measures of the probability of drawing genes identical by descent from all four combinations of parents.

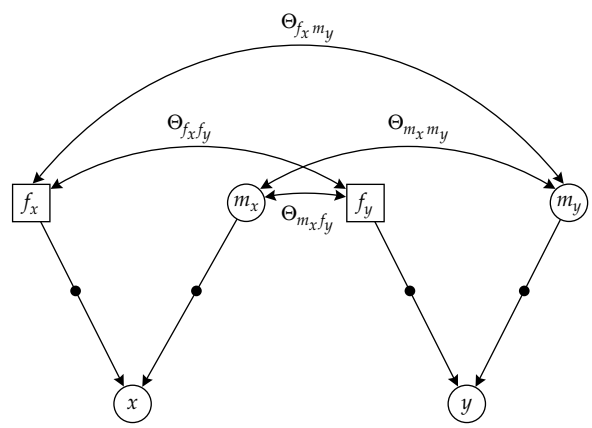


Figure 4.3 The analysis of the identity by descent of genotypes of individuals x and y. f_x and f_y represent fathers (which may be the same individual) of x and y, respectively, whereas m_x and m_y represent their mothers. Double-headed arrows between two parents represent coefficients of coancestry.

There are two ways by which the genotype of x can be identical by descent with that of y: (1) the gene descending from m_x may be identical by descent with that descending from m_y, and that from f_x identical by descent with that from f_y, or (2) the gene from m_x may be identical by descent with that from f_y, and that from f_x identical by descent with that from m_y. Thus, the coefficient of fraternity is defined as

$$\Delta_{xy} = \Theta_{m_x m_y}\Theta_{f_x f_y} + \Theta_{m_x f_y}\Theta_{f_x m_y} \tag{4.13}$$

Two examples will suffice to illustrate the use of this equation. First, consider the situation when x and y are full sibs, in which case the mothers are the same individual ($m_x = m_y = m$), as

are the fathers $f_x = f_y = f$. Equation 4.13 then reduces to

$$\Delta_{xy} = \Theta_{mm}\Theta_{ff} + \Theta_{mf}^2 \tag{4.14}$$

If the parents are unrelated, then $\Theta_{mf} = 0$; and if the parents are not inbred, then $\Theta_{mm} = \Theta_{ff} = 1/2$. Substituting these values into the above expression, we obtain $\Delta_{xy} = 1/4$.

Now consider the case of paternal half sibs, in which case the fathers are the same individual, but the mothers are different. Now,

$$\Delta_{xy} = \Theta_{m_x m_y}\Theta_{ff} + \Theta_{m_x f}\Theta_{f m_y} \tag{4.15}$$

Provided that the parents are unrelated, then $\Theta_{ff} = 1/2$ and $\Theta_{m_x m_y} = \Theta_{m_x f} = \Theta_{f m_y} = 0$, which yields $\Delta_{xy} = 0$. The genotypes of two individuals cannot be identical by descent if their maternally (or paternally) derived genes come from unrelated individuals.

Example 4.2. Returning to the figure in Example 4.1, what is Δ_{xy} for x = Royal Duke of Gloster and y = Princess Royal?

Designate the parents as f_x = Grand Duke of Gloster, m_x = Mimulus, f_y = Champion of England, and m_y = Carmine. Noting that Champion of England is the father of Grand Duke of Gloster and Mimulus, $\Theta_{f_x f_y} = \Theta_{m_x f_y} = (1/4)$. Counting the number of individuals in the paths of descent between the remaining two pairs of parents, $\Theta_{m_x m_y} = \Theta_{f_x m_y} = (1/2)^5$. Substituting into Equation 4.13, the probability that x and y have identical genotypes by descent at an arbitrary autosomal locus is

$$\Delta_{xy} = (1/4)(1/2)^5 + (1/2)^5(1/4) = (1/2)^6$$

We now have a complete system for describing the identity by descent at an arbitrary locus for any two individuals. For complex pedigrees, this can be a rather tedious process, but relatively simple algorithms exist for the computation of Θ_{xy} from simple information on parentage (Lynch and Walsh Chapter 26). The identity coefficients for several common relationships are summarized in Table 4.1.

Genetic Correlations for General Relationships

The above results for the contribution when relatives share one and two alleles IBD suggests the general expression for the covariance between (noninbred) relatives.

If r_{xy} = (1/2) Prob(relatives x and y have one allele IBD) + Prob(relatives x and y have both alleles IBD), and u_{xy} = Prob(relatives x and y have both alleles IBD), then the genetic covariance between x and y is given by

$$Cov(G_x, G_y) = r_{xy}V_A + u_{xy}V_D \tag{4.16}$$

It can be esaily shown that

$$2\Theta_{xy} = r_{xy}, \qquad u_{xy} = \Delta_{xy} \tag{4.17}$$

We can thus rewrite Equation 4.16 as

$$Cov(G_x, G_y) = 2\Theta_{xy}V_A + \Delta_{xy}V_D \tag{4.18}$$

If epistatic genetic variance is present, this can be generalized to

$$\begin{aligned} Cov(G_x, G_y) &= r_{xy}V_A + u_{xy}V_D + r_{xy}^2 V_{AA} + r_{xy}u_{xy}V_{AD} + u_{xy}^2 V_{DD} + \cdots \\ &= 2\Theta_{xy}V_A + \Delta_{xy}V_D + (2\Theta_{xy})^2 V_{AA} + 2\Theta_{xy}\Delta_{xy}V_{AD} + \Delta_{xy}^2 V_{DD} + \cdots \end{aligned} \tag{4.19}$$

Table 4.1 Identity coefficients for common relationships under the assumption of no inbreeding, in which case Δ_1 to $\Delta_6 = 0$.

Relationship	Θ_{xy}	Δ_{xy}
Parent–offspring	$\frac{1}{4}$	0
Grandparent–grandchild	$\frac{1}{8}$	0
Great grandparent–great grandchild	$\frac{1}{16}$	0
Half sibs	$\frac{1}{8}$	0
Full sibs, dizygotic twins	$\frac{1}{4}$	$\frac{1}{4}$
Uncle(aunt)–nephew(neice)	$\frac{1}{8}$	0
First cousins	$\frac{1}{16}$	0
Double first cousins	$\frac{1}{8}$	$\frac{1}{16}$
Second cousins	$\frac{1}{64}$	0
Monozygotic twins (clonemates)	$\frac{1}{2}$	1

Lecture 4 Problems

1. Again consider the Booroola Locus (Lecture 3). Suppose the environmental variance in litter size is 0.5 and there are no common environmental effects due to maternal environment or rearing families together. Assuming the Booroola locus is the only source of genetic variance, compute the regressions or intra-class correlations of litter size between the following relatives:

a: Offspring and midparent

b: Half sibs

c: Full sibs

Do the calculations for (i) populations where freq(B) = 0.3 and (ii) populations where freq(B) = 0.8. (Helpful hint – might want to use results from the Problem set for Lecture 3 (problems 4 and 5), unless you want the extra practice!)

2. What is the correlation between a grandparent and its grandchild?

Solutions to Lecture 4 Problems

1. Recalling from Problem 5 from the Lecture 3 problem set that :

for freq$(B) = 0.3 : \sigma_A^2 = 0.167$, $\sigma_D^2 = 0.002$, $\sigma_G^2 = 0.169$. Hence $\sigma_P^2 = 0.669$

for freq$(B) = 0.8 : \sigma_A^2 = 0.090$, $\sigma_D^2 = 0.001$, $\sigma_G^2 = 0.091$. Hence $\sigma_P^2 = 0.591$

a) Midparent-offspring regression: $b = V_A/V_p = 0.167/0.669 = 0.250$ for freq$(B) = 0.3$; 0.152 for freq$(B) = 0.8$.

b) Half-sib correlation: $t_{HS} = (1/4)V_A/V_P = 0.062$ for freq$(B) = 0.3$; 0.038 for freq$(B) = 0.8$.

c) Full-sib correlation: $t_{FS} = (V_A/2 + V_D/4)/V_P = 0.126$, for freq$(B) = 0.3$; 0.077 for freq$(B) = 0.8$.

2. A parent passes along a single allele to its offspring, and hence there is a $1/2$ change that the offspring passes on that allele to its offspring. Hence, $\Pr(2 \text{ alleles IBD}) = 0$, while $\Pr(1 \text{ alleles IBD}) = \Pr(0 \text{ alleles IBD}) = 1/2$. Hence, Cov(Grandparent, grandchild) $= \sigma_A^2/4$